

21

SEQ ID No. 1

Sequenced molecule: genomic DNA

Name: GBSS gene fragment from potato

Length of sequence: 342 bp

TGCATGTTTC	CCTACATTCT	ATTAGAAATC	GFGTTGTGGT	GTATAAAACGT	50
TGTTTCATAT	CTCATCTCAT	CTATTCTGAT	TTTGATTCTC	TTGCCTACTG	100
TAATCGGTGA	TAAATGTGAA	TGCTTCCTTT	CTTCTCAGAA	ATCAATTCT	150
GTTTTGTTTT	TGTTTCATCTG	TAGCTTATTC	TCTGGTAGAT	TCCCCTTTTT	200
GTAGACCACA	CATCAC	ATG	GCA	AGC	243
	Met	Ala	Ser	Ile	
	1			5	
TTT	GTG	TCA	AGA	AGC	
Phe	Val	Ser	Arg	Ser	
10				15	
CAA	ACT	TCA	CTA	GAC	
Asn	Thr	Ser	Leu	Asp	
				20	
ACC	AAA	TCA	ACC		
					285
TTG	TCA	CAG	ATA	GGA	
Leu	Ser	Gln	Ile	Gly	
25				30	
AGG	AAC	CAT	ACT	CTG	
Arg	Asn	His	Thr	Leu	
				35	
ACT	CAC	AAT			
					327
GGT	TTA	AGG	GCT	GTT	
Gly	Leu	Arg	Ala	Val	
				40	
					342

Sub B.

SEQ ID No. 2

Sequenced molecule: genomic DNA

Name: GBSS gene fragment from potato

Length of sequence: 2549 bp

AAC AAG CTT GAT GGG CTC CAA TCA ACA ACT AAT ACT AAG GTA	42
Asn Lys Leu Asp Gly Leu Gln Ser Thr Thr Asn Thr Lys Val	
45 50 55	
ACA CCC AAG ATG GCA TCC AGA ACT GAG ACC AAG AGA CCT GGA	84
Thr Pro Lys Met Ala Ser Arg Thr Glu Thr Lys Arg Pro Gly	
60 65 70	
TGC TCA GCT ACC ATT GTT TGT GGA AAG GGA ATG AAC TTG ATC	126
Cys Ser Ala Thr Ile Val Cys Gly Lys Gly Met Asn Leu Ile	
75 80	
TTT GTG GGT ACT GAG GTT GGT CCT TGG AGC AAA ACT GGT GGA	168
Phe Val Gly Thr Glu Val Gly Pro Trp Ser Lys Thr Gly Gly	
85 90 95	
CTA GGT GAT GTT CTT GGT GGA CTA CCA CCA GCC CTT GCA	207
Leu Gly Asp Val Leu Gly Gly Leu Pro Pro Ala Leu Ala	
100 105 110	
GTAAGTCITT CTTTCATTTG GTTACCTACT CATTCAATTAC TTATTTTGTT	257
TAGTTAGITT CTACTGCATC AGTCTTTTTA TCATTTAG GCC CGC GGA	304
Ala Arg Gly	
CAT CGG GTA ATG ACA ATA TCC CCC CGT TAT GAC CAA TAC AAA	346
His Arg Val Met Thr Ile Ser Pro Arg Tyr Asp Gln Tyr Lys	
115 120 125	
GAT GCT TGG GAT ACT GGC GTT GCG GTT GAG GTACATCTTC	386
Asp Ala Trp Asp Thr Gly Val Ala Val Glu	
130 135	
CTATATTGAT ACGGTACAAT ATTGTTCTCT TACATTTCTT GATTCAAGAA	436
TGTGATCATC TGCAG GTC AAA GTT GGA GAC AGC ATT GAA ATT GTT	481
Val Lys Val Gly Asp Ser Ile Glu Ile Val	
140 145	
CGT TTC TTT CAC TGC TAT AAA CGT GGG GTT GAT CGT GTT TTT	523
Arg Phe Phe His Cys Tyr Lys Arg Gly Val Asp Arg Val Phe	
150 155 160	
GTT GAC CAC CCA ATG TTC TTG GAG AAA GTAAGCATAT	560
Val Asp His Pro Met Phe Leu Glu Lys	
165 170	

Sub B1

TATGATTATG AATCCGTCCT GAGGGATACG CAGAACAGGT CATTTTGAGT 610
 ATCTTTTAAC TCTACTGGTG CTTTACTCT TTTAAG GTT TGG GGC AAA 658
 Val Trp Gly Lys 175

 ACT GGT TCA AAA ATC TAT GGC CCC AAA GCT GGA CTA GAT TAT 700
 Thr Gly Ser Lys Ile Tyr Gly Pro Lys Ala Gly Leu Asp Tyr 180 185

 CTG GAC AAT GAA CTT AGG TTC AGC TTG TTG TGT CAA 736
 Leu Asp Asn Glu Leu Arg Phe Ser Leu Leu Cys Gln 190 195 200

 GTAAGTTAGT TACTCTTGAT TTTTATGTGG CATTTTACTC TTTTGTCTTT 786
 AATCGTTTTT TTAACCTTGT TTTCTCAG GCA GCC CTA GAG GCA CCT 832
 Ala Ala Leu Glu Ala Pro 205

 AAA GTT TTG AAT TTG AAC AGT AGC AAC TAC TTC TCA GGA CCA 874
 Lys Val Leu Asn Leu Asn Ser Ser Asn Tyr Phe Ser Gly Pro 210 215 220

 TAT G GTAATTAACA CATCTAGTT TCAGAAACT CCTTACTATA 918
 Tyr G

 TCATTGTAGG TAATCATCTT TATTTTGCCT ATTCCTGCAG GA GAG GAT 966
 ly Glu Asp 225

 GTT CTC TTC ATT GCC AAT GAT TGG CAC ACA GCT CTC ATT CCT 1008
 Val Leu Phe Ile Ala Asn Asp Trp His Thr Ala Leu Ile Pro 230 235

 TGC TAC TTG AAG TCA ATG TAC CAG TCC AGA GGA ATC TAC TTG 1050
 Cys Tyr Leu Lys Ser Met Tyr Gln Ser Arg Gly Ile Tyr Leu 240 245 250

 AAT GCC AAG GTAAAATTTC TTTGTATTCA CTCGATTGCA 1089
 Asn Ala Lys 255

 CGTTACCCTG CAAATCAGTA AGGTTGTATT AATATATGAT AAATTTTCACA 1139
 TTGCCTCCAG GTT GCT TTC TGC ATC CAT AAC ATT GCC TAC CAA 1182
 Val Ala Phe Cys Ile His Asn Ile Ala Tyr Gln 260 265

 GGT CGA TTT TCT TTC TCT GAC TTC CCT CTT CTC AAT CTT CCT 1224
 Gly Arg Phe Ser Phe Ser Asp Phe Pro Leu Leu Asn Leu Pro 270 275 280

 GAT GAA TTC AGG GGT TCT TTT GAT TTC ATT GAT GGG TAT 1263
 Asp Glu Phe Arg Gly Ser Phe Asp Phe Ile Asp Gly Tyr 285 290

 GTATTATGC TTGAAATCAG AACTCCAACT TTTGAACTC TTTTGATGCT 1313

AGTAAATTGA GTTTTAAAA TTTTGCAGAT ATGAG AAG CCT GTT AAG 1360
 Lys Pro Val Lys
 295

GGT AGG AAA ATC AAC TGG ATG AAG GCT GGG ATA TTA GAA TCA 1402
 Gly Arg Lys Ile Asn Trp Met Lys Ala Gly Ile Leu Glu Ser
 300 305 310

CAT AGG GTG GTT ACA GTG AGC CCA TAC TAT GCC CAA GAA CTT 1444
 His Arg Val Val Thr Val Ser Pro Tyr Tyr Ala Gln Glu Leu
 315 320 325

GTC TCT GCT GTT GAC AAG GGA GTT GAA TTG GAC AGT GTC CTT 1486
 Val Ser Ala Val Asp Lys Gly Val Glu Leu Asp Ser Val Leu
 330 335 340

CGT AAG ACT TGC ATA ACT GGG ATT GTG AAT GGC ATG GAT ACA 1528
 Arg Lys Thr Cys Ile Thr Gly Ile Val Asn Gly Met Asp Thr
 345 350

CAA GAG TGG AAC CCA GCG ACT GAC AAA TAC ACA GAT GTC AAA 1570
 Gln Glu Trp Asn Pro Ala Thr Asp Lys Tyr Thr Asp Val Lys
 355 360 365

TAC GAT ATA ACC ACT GTAAGATAAG ATTTTCCGA CTCCAGTATA 1615
 Tyr Asp Ile Thr Thr
 370

TACTAAATTA TTTTGTATGT TTATGAAATT AAAGAGTTCT TGCTAATCAA 1665
 AATCTCTATA CAG GTC ATG GAC GCA AAA CCT TTA CTA AAG GAG 1708
 Val Met Asp Ala Lys Pro Leu Leu Lys Glu
 375 380

GCT CTT CAA GCA GCA GTT GGC TTG CCT GTT GAC AAG AAG ATC 1756
 Ala Leu Gln Ala Ala Val Gly Leu Pro Val Asp Lys Lys Ile
 385 390 395

CCT TTG ATT GGC TTC ATC GGC AGA CTT GAG GAG CAG AAA GGT 1792
 Pro Leu Ile Gly Phe Ile Gly Arg Leu Glu Glu Gln Lys Gly
 400 405 410

TCA GAT ATT CTT GTT GCT GCA ATT CAC AAG TTC ATC GGA TTG 1834
 Ser Asp Ile Leu Ala Val Ala Ile His Lys Phe Ile Gly Leu
 415 420 425

GAT GTT CAA ATT GTA GTC CTT GTAAGTACCA AATGGACTCA 1875
 Asp Val Gln Ile Val Val Leu
 430

TGGTATCTCT CTTGTTGAGT TTACTTGTGC CGAAACTGAA ATTGACCTGC 1925
 TACTCATCTT ATGCATCAG GGA ACT GGC AAA AAG GAG TTT GAG 1968
 Gly Thr Gly Lys Lys Glu Phe Glu
 435 440

Sub
B

25

CAG GAG ATT GAA CAG CTC GAA GTG TTG TAC CCT AAC AAA GCT	2010
Gln Glu Ile Glu Gln Leu Glu Val Leu Tyr Pro Asn Lys Ala	
445 450	
AAA GGA GTG GCA AAA TTC AAT GTC CCT TTG GCT CAC ATG ATC	2052
Lys Gly Val Ala Lys Phe Asn Val Pro Leu Ala His Met Ile	
455 460 465	
ACT GCT GGT GCT GAT TTT ATG TTG GTT CCA AGC AGA TTT GAA	2094
Thr Ala Gly Ala Asp Phe Met Leu Val Pro Ser Arg Phe Glu	
470 475 480	
CCT TGT GGT CTC ATT CAG TTA CAT GCT ATG CGA TAT GGA ACA	2136
Pro Cys Gly Leu Ile Gln Leu His Ala Met Arg Tyr Gly Thr	
485 490 495	
GTAAGAACCA GAAGAGCTTG TACCTTTTTC CTGAGTTTTT AAAAAAAGAA	2186
TCATAAGACC TTGTTTTTCCA TCTAAAGTTT AATAACCAAC TAAATGTTAC	2236
TGCAGCAAGC TTTTCATTTT TGAAAATTGG TTATCTGATT TTAACGTAAT	2286
CACATGTGAG TCAG GTA CCA ATC TGT GCA TCG ACT GGT GGA CTT	2330
Val Pro Ile Cys Ala Ser Thr Gly Gly Leu	
500 505	
GTT GAC ACT GTG AAA GAA GGC TAT ACT GGA TTC CAT ATG GGA	2372
Val Asp Thr Val Lys Glu Gly Tyr Thr Gly Phe His Met Gly	
510 515 520	
GCC TTC AAT GTT GAA GTATGTGATT TTACATCAAT TGTGTACTTG	2417
Ala Phe Asn Val Glu	
525	
TACATGGTCC ATTCTCGTCT TGATATACCC CTTGTTGCAT AAACATTAAC	2467
TTATTGCTTC TTGAATTTGG TTAG TGC GAT GTT GTT GAC CCA GCT	2512
Cys Asp Val Val Asp Pro Ala	
530	
GAT GTG CTT AAG ATA GTA ACA ACA GTT GCT AGA GCT C	2549
Asp Val Leu Lys Ile Val Thr Thr Val Ala Arg Ala	
535 540	

Sub B.1

26

SEQ ID No. 3

Sequenced molecule: genomic DNA

Name: GBSS gene fragment from potato

Length of sequence: 492 bp

GAG CTC TCC TGG AAG	GTAAGTGTGA ATTTGATAAT TTGCGTAGGT	45
Glu Leu Ser Trp Lys		
565		
ACTTCAGTTT GTTGTTCG TCA	GCACTGA TGGATTCCAA CTGGTGTTC	95
TGCAG	GAA CCT GCC AAG AAA TGG GAG ACA TTG	127
	Glu Pro Ala Lys Lys Trp Glu Thr Leu	
	570 575	
CTA TTG GGC TTA GGA GCT TCT GGC AGT GAA CCC GGT GTT GAA	169	
Leu Leu Gly Leu Gly Ala Ser Gly Ser Glu Pro Gly Val Glu		
580 585 590		
GGG GAA GAA ATC GCT CCA CTT GCC AAG GAA AAT GTA GCC ACT	211	
Gly Glu Glu Ile Ala Pro Leu Ala Lys Glu Asn Val Ala Thr		
595 600 605		
CCT TAA	ATGAGCTTTG GTTATCCTTG TTTCAACAAT AAGATCATTA	257
Pro ***		
606		
AGCAAACGTA TTTACTAGCG AACTATGTAG AACCCCTATTA TGGGGTCTCA	307	
ATCATCTACA AAATGATTGG TTTTGTCTGG GGAGCAGCAG CATATAAGGC	357	
TGTAAATCC TGGTTAATGT TTTTGTAGGT AAGGGCTATT TAAGGTGGTG	407	
TGGATCAAG TCAATAGAAA ATAGTTATTA CTAACGTTTG CAACTAAATA	457	
CTTAGTATG TAGCATAAAT AATACTAGAA CTAGT	492	

Sub B.

21

27

SEQ ID No. 4

Sequenced molecule: genomic DNA

Name: Promoter for the GBSS gene from potato

Length of sequence: 987 bp

AAGCTTTAAC	GAGATAGAAA	ATTATGTTAC	TCCGTTTTGT	TCATTACTTA	50
ACAAATGCAA	CAGTATCTTG	TACCAAATCC	TTTCTCTCTT	TTCAAACCTT	100
TCTATTTGGC	TGTTGACGGA	GTAATCAGGA	TACAAACCAC	AAGTATTTAA	150
TTGACTCCTC	CGCCAGATAT	TATGATTTAT	GAATCCTCGA	AAAGCCTATC	200
CATTAAGTCC	TCATCTATGG	ATATACTTGA	CAGTATCTTC	CTGTTTGGGT	250
ATTTTTTTTT	CCTGCCAAGT	GGAACGGAGA	CATGTTATGA	TGTATACGGG	300
AAGCTCGTTA	AAAAAAAAAT	CAATAGGAAG	AAATGTAACA	AACATTGAAT	350
GTTGTTTTTA	ACCATCCTTC	CTTTAGCAGT	GTATCAATTT	TGTAATAGAA	400
CCATGCATCT	CAATCTTAAT	ACTAAAATGC	AACTTAATAT	AGGCTAAACC	450
AAGATAAAGT	AATGTATTCA	ACCTTTAGAA	TTGTGCATTC	ATAATTAGAT	500
CTTGTTTGTC	GTAATAAATT	AGAAAATATA	TTTACAGTAA	TTTGGAATAC	550
AAAGCTAAGG	GGGAAGTAAC	TAATATTCTA	GTGGAGGGAG	GGACCAGTAC	600
CAGTACCTAG	ATATTATTTT	TAATTACTAT	AATAATAATT	TAATTAACAC	650
GAGACATAGG	AATGTCAAGT	GGTAGCGTAG	GAGGGAGTTG	GTTTAGTTTT	700
TTAGATACTA	GCAGACAGAA	CCGGACGGCC	CATTGCAAGG	CCAAGTTGAA	750
GTCCAGCCGT	GAATCAACAA	AGAGAGGGCC	CATAATACTG	TCGATGAGCA	800
TTTCCCTATA	ATACAGTGTC	CACAGTTGCC	TTCTGCTAAG	GGATAGCCAC	850
CCGCTATTCT	CTTGACACGT	GTCACGAAA	CCTGCTACAA	ATAAGGCAGG	900
CACCTCCTCA	TTCTCACTCA	CTCACTCACA	CAGCTCAACA	AGTGGTAACT	950
TTTACTCAAC	TCCTCCAATT	ATTTCTGATT	TCATGCA		987

Sub B1

SEQ ID No. 5

Sequenced molecule: genomic DNA

Name: GBSS gene from potato

Length of sequence: 4964 bp

AAGCTTTAAC	GAGATAGAAA	ATTATGTTAC	TCCGTTTTGT	TCATTACTTA	50
ACAAATGCAA	CAGTATCTTG	TACCAAATCC	TTTCTCTCTT	TTCAAACCTT	100
TCTATTTGGC	TGTTGACGGA	GTAATCAGGA	TACAAACCAC	AAGTATTTAA	150
TTGACTCCTC	CGCCAGATAT	TATGATTTAT	GAATCCTCGA	AAAGCCTATC	200
CATTAAGTCC	TCATCTATGG	ATATACTTGA	CAGTATCTTC	CTGTTTGGGT	250
ATTTTTTTTT	CCTGCCAAGT	GGAAAGGAGA	CATGTTATGA	TGTATACGGG	300
AAGCTCGTTA	AAAAAAAATA	CAATAGGAAG	AAATGTAACA	AACATTGAAT	350
GTTGTTTTTA	ACCATCCTTC	CTTTAGCAGT	GTATCAATTT	TGTAATAGAA	400
CCATGCATCT	CAATCTTAAT	ACTAAAATGC	AACCTTAATAT	AGGCTAAACC	450
AAGATAAAGT	AATGTATTCA	ACCTTTAGAA	TTGTGCATTC	ATAATTAGAT	500
CTTGTTTGTC	GTAAAAAATT	ACAAAATATA	TTTACAGTAA	TTTGAATAC	550
AAAGCTAAGG	GGGAAGTAAC	TAATATTCTA	GTGGAGGGAG	GGACCAGTAC	600
CAGTACCTAG	ATATTATFTT	TAATTACTAT	AATAATAATT	TAATTAACAC	650
GAGACATAGG	AATGTCAAGT	GGTAGCGTAG	GAGGGAGTTG	GTTTAGTTTT	700
TTAGATACTA	GGAGACAGAA	CCGGACGGCC	CATTGCAAGG	CCAAGTTGAA	750
GTCCAGCCGT	GAATCAACAA	AGAGAGGGCC	CATAATACTG	TCGATGAGCA	800
TTTCCCTATA	ATACAGTGTC	CACAGTTGCC	TTCTGCTAAG	GGATAGCCAC	850
CCGCTATTCT	CTTGACACCT	GTCACTGAAA	CCTGCTACAA	ATAAGGCAGG	900
CACCTCCTCA	TTCTCACTCA	CTCACTCACA	CAGCTCAACA	AGTGGTAACT	950
TTTACTCATC	TCCTCCAATT	ATTTCTGATT	TCATGCATGT	TTCCCTACAT	1000
TCTATTATGA	ATCGTGTTGT	GGTGTATAAA	CGTTGTTTCA	TATCTCATCT	1050
CATCTATTCT	GATTTTGATT	CTCTTGCTTA	CTGTAATCGG	TGATAAATGT	1100
GAATGCTTCC	TTTCTTCTCA	GAAATCAATT	TCTGTTTTGT	TTTTGTTCAT	1150
CTGTAGCTTA	TTCTCTGGTA	GATTCCCCTT	TTTGTAGACC	ACACATCAC	1199
ATG GCA AGC ATC ACA GCT TCA CAC CAC TTT GTG TCA AGA AGC					1241
Met Ala Ser Ile Thr Ala Ser His His Phe Val Ser Arg Ser					
1	5		10		
CAA ACT TCA CTA GAC ACC AAA TCA ACC TTG TCA CAG ATA GGA					1283
Gln Thr Ser Leu Asp Thr Lys Ser Thr Leu Ser Gln Ile Gly					
15	20		25		
CTC AGG AAG CAT ACT CTG ACT CAC AAT GGT TTA AGG GCT GTT					1325
Leu Arg Asn His Thr Leu Thr His Asn Gly Leu Arg Ala Val					
30	35		40		
AAC AAG CTT GAT GGG CTC CAA TCA ACA ACT AAT ACT AAG GTA					1367
Asn Lys Leu Asp Gly Leu Gln Ser Thr Thr Asn Thr Lys Val					
45	50		55		
ACA CCC AAG ATG GCA TCC AGA ACT GAG ACC AAG AGA CCT GGA					1409
Thr Pro Lys Met Ala Ser Arg Thr Glu Thr Lys Arg Pro Gly					
60	65		70		
TGC TCA GCT ACC ATT GTT TGT GGA AAG GGA ATG AAC TTG ATC					1451
Cys Ser Ala Thr Ile Val Cys Gly Lys Gly Met Asn Leu Ile					
75	80				
TTT GTG GGT ACT GAG GTT GGT CCT TGG AGC AAA ACT GGT GGA					1493
Phe Val Gly Thr Glu Val Gly Pro Trp Ser Lys Thr Gly Gly					
85	90		95		

CTA GGT GAT GTT CTT GGT GGA CTA CCA CCA GCC CTT GCA 1532
 Leu Gly Asp Val Leu Gly Gly Leu Pro Pro Ala Leu Ala
 100 105 110

GTAAGTCTTT CTTTCATTG GTTACCTACT CATTCAATTAC TTATTTTGTG 1582
 TAGTTAGTTT CTTACTGCATC AGTCTTTTAA TCATTTAG GCC CGC GGA 1629
 Ala Arg Gly

CAT CGG GTA ATG ACA ATA TCC CCC CGT TAT GAC CAA TAC AAA 1671
 His Arg Val Met Thr Ile Ser Pro Arg Tyr Asp Gln Tyr Lys
 115 120 125

GAT GCT TGG GAT ACT GGC GTT GCG GTT GAG GTACATCTTC 1711
 Asp Ala Trp Asp Thr Gly Val Ala Val Glu
 130 135

CTATATTGAT ACGGTACAAT ATTGTTCTCT TACATTTCTT GATTCAAGAA 1761
 TGTGATCATC TGCAG GTC AAA GTT GGA GAC AGC ATT GAA ATT GTT 1806
 Val Lys Val Gly Asp Ser Ile Glu Ile Val
 140 145

CGT TTC TTT CAC TGC TAT AAA CGT GGG GTT GAT CGT GTT TTT 1848
 Arg Phe Phe His Cys Tyr Lys Arg Gly Val Asp Arg Val Phe
 150 155 160

GTT GAC CAC CCA ATG TTC TTG GAG AAA GTAAGCATAT 1885
 Val Asp His Pro Met Phe Leu Glu Lys
 165 170

TATGATTATG AATCGGTCCT GAGGGATACG CAGAACAGGT CATTTTGAGT 1935
 ATCTTTTAAC TCTACTGGTG CTTTACTCT TTTAAG GTT TGG GGC AAA 1983
 Val Trp Gly Lys
 175

ACT GGT TCA AAA ATC TAT GGC CCC AAA GCT GGA CTA GAT TAT 2025
 Thr Gly Ser Lys Ile Tyr Gly Pro Lys Ala Gly Leu Asp Tyr
 180 185

CTG GAC AAT GAA CTT AGG TTC AGC TTG TTG TGT CAA 2061
 Leu Asp Asn Glu Leu Arg Phe Ser Leu Leu Cys Gln
 190 195 200

GTAAGTTAGT TACTCTTGAT TTTTATGTGG CATTTTACTC TTTTGTCTTT 2111
 AATCGTTTTT TTAACCTTGT TTTCTCAG GCA GCC CTA GAG GCA CCT 2157
 Ala Ala Leu Glu Ala Pro
 205

AAA GTT TTG AAT TTG AAC AGT AGC AAC TAC TTC TCA GGA CCA 2199
 Lys Val Leu Asn Leu Asn Ser Ser Asn Tyr Phe Ser Gly Pro
 210 215 220

TAT G GTAATTAACA CATCCTAGTT TCAGAAACT CCTTACTATA 2243
Tyr G

TCATTGTAGG TAATCATCTT TATTTTGCCT ATTCTCTGCAG GA GAG GAT 2291
ly Glu Asp
225

GTT CTC TTC ATT GCC AAT GAT TGG CAC ACA GCT CTC ATT CCT 2333
Val Leu Phe Ile Ala Asn Asp Trp His Thr Ala Leu Ile Pro
230 235

TGC TAC TTG AAG TCA ATG TAC / CAG TCC AGA GGA ATC TAC TTG 2375
Cys Tyr Leu Lys Ser Met Tyr / Gln Ser Arg Gly Ile Tyr Leu
240 245 250

AAT GCC AAG GTAAATTTC TTTGTATTCA CTCGATTGCA 2414
Asn Ala Lys
255

CGTTACCCCTG	CAAATCAGTA	AGGTTGTATT	AATATATGAT	AAATTTTACA	2464
TTGCCCTCCAG	GTT GCT TTC	TGC ATC CAT	AAC ATT GCC	TAC CAA	2507
	Val Ala Phe	Cys Ile His	Asn Ile Ala	Tyr Gln	
	260		265		

GGT CGA TTT TCT TTC TCT GAC TTC CCT CTT CTC AAT CTT CCT 2549
Gly Arg Phe Ser Phe Ser Asp Phe Pro Leu Leu Asn Leu Pro
270 275 280

GAT GAA TTC AGG GGT TCT TTT GAT TTC ATT GAT GGG TAT 2588
Asp Glu Phe Arg Gly Ser Phe Asp Phe Ile Asp Gly Tyr
285 290

GTATTTATGC	TTGAAATCAG	ACCTCCAAC	TTTGAAGCTC	TTTGTATGCT	2638
AGTAAATGGA	GTTTTTAAAA	TTTTGCAGAT	ATGAG AAG	CCT GTT AAG	2685
			Lys	Pro Val Lys	
			295		

GGT AGG AAA/ATC AAC TGG ATG AAG GCT GGG ATA TTA GAA TCA 2727
Gly Arg Lys Ile Asn Trp Met Lys Ala Gly Ile Leu Glu Ser
300 305 310

CAT	AGG	GTG	GTT	ACA	GTG	AGC	CCA	TAC	TAT	GCC	CAA	GAA	CTT	2769
His	Arg	Val	Val	Thr	Val	Ser	Pro	Tyr	Tyr	Ala	Gln	Glu	Leu	
		315					320					325		

GTC TCT GCT GTT GAC AAG GGA GTT GAA TTG GAC AGT GTC CTT 2811
Val Ser Ala Val Asp Lys Gly Val Glu Leu Asp Ser Val Leu
330 335 340

CGT AAG ACG GCG ATA ACG GGG ATT GTG AAT GGC ATG GAT ACA 2853
 Arg Lys Thr Cys Ile Thr Gly Ile Val Asp Gly Met Asp Thr
 345 350

Sub B.

5

CAA GAG TGG AAC CCA GCG ACT GAC AAA TAC ACA GAT GTC AAA 2895
 Gln Glu Trp Asn Pro Ala Thr Asp Lys Tyr Thr Asp Val Lys
 355 360 365

TAC GAT ATA ACC ACT GTAAGATAAG ATTTTCCGA CTCCAGTATA 2940
 Tyr Asp Ile Thr Thr
 370

TACTAAATTA TTTTGTATGT TTATGAAATT AAAGAGTTCT TGCTAATCAA 2990
 AATCTCTATA CAG GTC ATG GAC GCA AAA CCT TTA CTA AAG GAG 3033
 Val Met Asp Ala Lys Pro Leu Leu Lys Glu
 375 380

GCT CTT CAA GCA GCA GTT GGC TTG CCT GTT GAC AAG AAG ATC 3075
 Ala Leu Gln Ala Ala Val Gly Leu Pro Val Asp Lys Lys Ile
 385 390 395

CCT TTG ATT GGC TTC ATC GGC AGA CTT GAG GAG CAG AAA GGT 3117
 Pro Leu Ile Gly Phe Ile Gly Arg Leu Glu Glu Gln Lys Gly
 400 405 410

TCA GAT ATT CTT GTT GCT GCA ATT CAC AAG TTC ATC GGA TTG 3159
 Ser Asp Ile Leu Ala Val Ala Ile His Lys Phe Ile Gly Leu
 415 420 425

GAT GTT CAA ATT GTA GTC CTT GTAAGTACCA AATGGACTCA 3200
 Asp Val Gln Ile Val Val Leu
 430

TGGTATCTCT CTTGTTGAGT TTAAGTGTGC CGAACTGAA ATTGACCTGC 3250
 TACTCATCCT ATGCATCAG GGA ACT GGC AAA AAG GAG TTT GAG 3293
 Gly Thr Gly Lys Lys Glu Phe Glu
 435 440

CAG GAG ATT GAA CAG CTC GAA GTG TTG TAC CCT AAC AAA GCT 3335
 Gln Glu Ile Glu Gln Leu Glu Val Leu Tyr Pro Asn Lys Ala
 445 450

AAA GGA GTG GCA AAA TTC AAT GTC CCT TTG GCT CAC ATG ATC 3377
 Lys Gly Val Ala Lys Phe Asn Val Pro Leu Ala His Met Ile
 455 460 465

ACT GCT GGT GCT GAT TTT ATG TTG GTT CCA AGC AGA TTT GAA 3419
 Thr Ala Gly Ala Asp Phe Met Leu Val Pro Ser Arg Phe Glu
 470 475 480

CCT TGT GGT CTC ATT CAG TTA CAT GCT ATG CGA TAT GGA ACA 3461
 Pro Cys Gly Leu Ile Gln Leu His Ala Met Arg Tyr Gly Thr
 485 490 495

GTAAGAACCA GAAGAGCTTG TACCTTTTCTA CTGAGTTTTT AAAAAAGAA 3511
 TCATAGACC TTGTTTTCCA TCTAAAGTTT AATAACCAAC TAAATGTTAC 3561
 TGCAGCAAGC TTTTCATTTT TGAATATTGG TTATCTGATT TTAACGTAAT 3611

Sub B.

CACATGTGAG TCAG GTA CCA ATC TGT GCA TCG ACT GGT GGA CTT	3655
Val Pro Ile Cys Ala Ser Thr Gly Gly Leu	
500 505	
GTT GAC ACT GTG AAA GAA GGC TAT ACT GGA TTC CAT ATG GGA	3697
Val Asp Thr Val Lys Glu Gly Tyr Thr Gly Phe His Met Gly	
510 515 520	
GCC TTC AAT GTT GAA GTATGTGATT TTACATCAAT TGTGTACTTG	3742
Ala Phe Asn Val Glu	
525	
TACATGGTCC ATTCTCGTCT TGATATACCC CTTGTTGCAT AAACATTAAC	3792
TTATTGCTTC TTGAATTTGG TTAG TGC GAT GTT GTT GAC CCA GCT	3837
Cys Asp Val Val Asp Pro Ala	
530	
GAT GTG CTT AAG ATA GTA ACA ACA GTT GCT AGA GCT CTT GCA	3879
Asp Val Leu Lys Ile Val Thr Thr Val Ala Arg Ala Leu Ala	
535 540 545	
GTC TAT GGC ACC CTC GCA TTT GCT GAG ATG ATA AAA AAT TGC	3921
Val Tyr Gly Thr Leu Ala Phe Ala Glu Met Ile Lys Asn Cys	
550 555 560	
ATG TCA GAG GAG CTC TCC TGG AAG GTAAGTGTGA ATTTGATAAT	3965
Met Ser Glu Glu Ile Ser Trp Lys	
565	
TTGCGTAGGT ACTTCAGTTT GTTGTCTCTG TCAGCACTGA TGGATTCCAA	4015
CTGGTGTTCT TGCAG GAA CCT GCC AAG AAA TGG GAG ACA TTG	4057
Glu Pro Ala Lys Lys Trp Glu Thr Leu	
570 575	
CTA TTG GGC TTA GGA GCT TCT GGC AGT GAA CCC GGT GTT GAA	4099
Leu Leu Gly Leu Gly Ala Ser Gly Ser Glu Pro Gly Val Glu	
580 585 590	
GGG GAA GAA ATC GCT CCA CTT GCC AAG GAA AAT GTA GCC ACT	4141
Gly Glu Glu Ile Ala Pro Leu Ala Lys Glu Asn Val Ala Thr	
595 600 605	
CCT TAA ATGAGCTTTG GTTATCCTTG TTTCAACAAT AAGATCATT	4187
Pro ***	
606	
AGCAACCGTA TTTACTAGCG AACTATGTAG AACCCCTATTA TGGGGTCTCA	4237
ATCATCTACA AAATGATTGG TTTTGTGCTGG GGAGCAGCAG CATATAAGGC	4287
TGTAAATCC TGGTTAATGT TTTTGTAGGT AAGGGCTATT TAAGGTGGTG	4337
TGGATCAAG TCAATAGAAA ATAGTTATTA CTAACGTTTG CAACTAAATA	4387
CTTAGTAATG TAGCATAAAT AATACTAGAA CTAGTAGCTA ATATATATGC	4437
GTCAATTTGT TGTACCTTTT CTTCATAAAT TATTTGCAGT ACATATATAA	4487
TGAAATTTAC CCAAGGAATC AATGTTTCTT GCTCCGTCCT CCTTTGATGA	4537
TTTTTTACCC AATACAGAGC TAGTGTGTTA TGTATATAAT TTTGTTTAA	4587

Sub B.

AGAAGTAATC	AAATTCAAAT	TAGTTGTTTG	GTCATATGAA	AGAAGCTGCC	4637
AGGCTAACTT	TGAGGAGATG	GCTATTGAAT	TTCAAAATGA	TTATGTGAAA	4687
ACAATGCAAC	ATCTATGTCA	ATCAACACTT	AAATTATTGC	ATTTAGAAAG	4737
ATATTTTTGA	GCCCATGACA	CATTCAATCA	TAAAGTAAGG	TAGTATGTAT	4787
GATTGAATGG	ACTACAGCTC	AATCAAAGCA	TCTCCTTTAC	ATAACGGCAC	4837
TGTCTCTTGT	CTACTACTCT	ATTGGTAGTA	GTAGTAGTAA	TTTTACAATC	4887
CAAATTGAAT	AGTAATAAGA	TGCTCTCTAT	TTACTAAAGT	AGTAGTATTA	4937
TTCTTTCGTT	ACTCTAAAGC	AACAAAA			4964

Sub B₁